Campus de l'Université de Montréal
C.P. 6128, succursale Centre ville
Montréal QC H3C 3J7 Canada

Telephone (514) 343-7575
Télécopieur (514) 343-7121

crt@crt.umontreal.ca
www.crt.umontreal.ca

CRT-2002-32

Centre de recherche sur les transports (C.R.T.)
Centre for Research on Transportation

AN ANALYTICAL APPROACH TO THE FACILITY LOCATION AND CAPACITY ACQUISITION PROBLEM UNDER DEMAND UNCERTAINTY

by

Abdullah Dasci
Gilbert Laporte

October 2002

103-06368/

PUBLICATION
CRT-2002-32

AN ANALYTICAL APPROACH TO THE FACILITY LOCATION AND CAPACITY ACQUISITION PROBLEM UNDER DEMAND UNCERTAINTY*

by

Abdullah Dasci[1]
Gilbert Laporte[1]

* This research was supported in part by the Canada Research Chair in Distribution Management and by the Natural Sciences and Engineering Research Council of Canada (NSERC).

[1] Canada Research Chair in Distribution Management, HEC Montréal, 3000 chemin de la Côte-Sainte-Catherine, Montréal H3T 2A7 Canada and Centre de recherche sur les transports, Université de Montréal, C.P. 6128, succursale Centre-Ville, Montréal H3C 3J7 Canada (dasci@crt.umontreal.ca, gilbert@crt.umontreal.ca)

Dépôt légal – Bibliothèque nationale du Québec
Bibliothèque nationale du Canada, 2002

Abstract

This article presents an analysis of facility location and capacity acquisition under demand uncertainty. A novel methodology is proposed, in which the focus is shifted from the precise representation of facility locations to the market areas they serve. This is an extension of the optimal market area approach in which market area size and facility capacity are determined to minimize the total cost associated with fixed facility opening, variable capacity acquisition, transportation, and shortage. The problem has two variants depending on whether the firm satisfies shortages by outsourcing or shortages become lost sales. The analytical approach simplifies the problem considerably and leads to intuitive and insightful models. Among several other results, it is shown that fewer facilities are set up under lost sales than under outsourcing. It is also shown that the total cost in both models is relatively insensitive to small deviations in optimal capacity choices and parameter estimations.

Keywords: location, distribution.

Résumé

Dans cet article, on présente une analyse d'un problème de localisation d'installation avec coût d'acquisition de capacité dans un contexte de demande aléatoire. On propose une nouvelle méthodologie qui met l'emphase sur l'aire de marché desservie par l'installation par opposition au choix précis de sa localisation. Cette étude constitue une extension du modèle d'aires de marché dans lequel on détermine des aires de marché dans le but de minimiser un coût total comprenant des coûts fixes d'ouverture d'installations, des coûts d'acquisition de capacité ainsi que des coûts de transport et de pénurie. Le problème comporte deux variantes définies selon que les pénuries sont satisfaites par d'autres installations ou qu'elles se traduisent en ventes perdues. L'approche analytique préconisée simplifie le problème de manière significative et donne lieu à des interprétations intuitives. On montre entre autres que la seconde variante donne lieu à l'ouverture d'un plus grand nombre d'installations que la première variante. On montre de plus que le coût total dans les deux variantes est assez peu sensible à de petites variations de choix de capacité et aux estimations des paramètres du modèle.

Mots clés : localisation, distribution.

Introduction

Facility location and capacity acquisition are of vital importance to firms in capital intensive industries. These investments require large capital outlays and are irreversible in the short run. Since firms recognize that they have to live with their decisions for many years they resort to critical strategic planning. The multi-location capacity acquisition problem deals primarily with the issues of where to set up facilities, how much capacity to acquire, and which part of the market each facility should serve.

There are essentially two approaches to such problems. The more popular is the use of detailed mathematical programming models to determine the optimal location and size of facilities, as well as market areas. The second approach is a simplified analysis that focuses on the issues of facility size and market area under somewhat restrictive assumptions about topology and demand density, leaving the decision of precise locations for subsequent analyses. This approach, although less popular, is also well founded in classical location theory as it deals with the fundamental tradeoffs involved in these decisions. These models are called market area models (Erlenkotter[1]).

As far as system design is concerned, mathematical programs are more accurate and flexible than market area models. Mathematical programs allow the incorporation of multiple products, multiple echelons, temporal dynamics, and special constraints in a natural way, albeit at the cost of significant data and computational requirements. Also, their stability and accuracy diminish as the models become more complex. Moreover, while mathematical programs usually prescribe optimal solutions, they fail to explain *why* an optimal solution is what it is. Sensitivity analyses and the incorporation of uncertainty can also lead to countless computational requirements. Market area models, on the other hand, require less data and result in intuitive and insightful closed form solutions. Furthermore, they are less sensitive to errors in parameter estimation. However, they may not be well suited for complex systems. Thus, the two approaches do not really compete but rather complement each other. As observed by several researchers (Newell,[2] Geoffrion,[3,4] and Erlebacher and Meller[5]), market area models, which provide fundamental insights into the nature of the problem, can be used in conjunction with mathematical programming.

In this article, we present a market area model for the facility location and capacity acquisition problem under demand uncertainty. Most of the previous market area models are deterministic; therefore, the optimal size of a facility is equal to the total demand in the market area. A typical objective is to minimize the sum of fixed facility opening costs, variable acquisition and operating costs, and distribution costs. Under demand uncertainty, however, size becomes a decision to be optimized in the presence of potential excess capacity and shortage costs. Therefore, our model minimizes the total cost of facility opening, capacity acquisition, distribution, and shortage. This objective is essentially equivalent to revenue maximization.

Two streams of literature have produced a number of mathematical programming models for capacity planning and location. One of them, the dynamic multi-location capacity planning literature, starts with Manne's[6] seminal book on strategic planning for the cement, fertilizer, and caustic soda industries in India. Chapters 4, 12, and 13 present models under somewhat restrictive assumptions and propose heuristic solution procedures for those problems. Erlenkotter[7] has shown how such problems can be formulated as dynamic programs, but solutions can only be obtained for instances with a small number of locations. He has, therefore, resorted to a heuristic procedure termed "incomplete dynamic programming". More general versions of the problem have subsequently been developed by Rao and Rutenberg,[8] Fong and Srinivasan,[9–11] Klincewicz, Luss, and Yu[12] and Lim and Kim.[13] Even though all of the above models assume deterministic demand, no exact algorithm has been proposed for problems of realistic size. Recently, Ahmed and Garcia[14] have presented a two stage stochastic programming formulation for the dynamic capacity acquisition problem. In their model, all capacity decisions are made at the beginning of the planning period and only then are the customer assignments made, after uncertainty is resolved at each period. They have developed an exact solution method applicable to small size problems. The second stream of literature, dealing with facility location under uncertainty, is relatively less developed. Balachandran and Jain[15] have introduced a stochastic uncapacitated facility location problem with piecewise linear production cost. They have proposed a progressive linear approximation of the production costs, and then developed a branch-and-bound solution procedure for the resulting stochastic program with single recourse. LeBlanc[16] and Laporte, Louveaux, and Van hamme[17] have studied stochastic versions of capacitated facility location problems.

The first article proposes a heuristic based on solving a stochastic transportation problem, whereas the second one proposes a branch-and-cut based exact solution procedure. In all of these studies, with the exception of Ahmed and Garcia,[14] either the demand is deterministic but the models are dynamic, or the demand is stochastic but the models are static and capacity acquisition is not considered. Our model is also static and accommodates stochastic demand, but we consider capacity acquisition as a decision variable as well. To the best of our knowledge, there does not exist any previous analytical study on this problem.

We investigate two variants of the problem in which either the firm meets shortages by using an outside source, or shortages translate into lost sales. We refer to these cases as "Outsourcing" and "Lost Sales" models, respectively. Note that in the non-spatial production planning and inventory control models, which do not consider distribution, this distinction is not important. However, in our case it leads to different models and hence, different results. In contrast to production planning, the shortages we are concerned with are *not* those that result from short term demand fluctuations, which are assumed to be dealt separately by backlogging and inventory, but rather longer term chronic shortages. Hence, depending on future outsourcing options, either of the two variants could be appropriate. For example, if a chronic shortage in the industry was expected in the future, then probably the firm might not be able to satisfy shortages from outside and hence, the lost sales case would apply. On the other hand, if excess in the industry supply was anticipated, then outsourcing could be a viable option.

Although not all of our models can be solved in closed form, they nonetheless provide a number of important insights into the problem which may not be captured by mathematical programming. For example, we have shown that the lost sales case results in larger optimal market areas, i.e., in fewer facilities. Furthermore, we have shown that in the outsourcing case, the optimal market area only depends on expected demand, fixed costs, and expected transportation costs. In the lost sales case, however, it also depends on demand distribution and shortage cost. We have also observed that in both models, the total costs are relatively flat around the optimal capacity and thus insensitive to small changes in model parameters and optimal capacity. However, they become more sensitive as the coefficient of variation of demand decreases or as unit shortage cost increases or as variable capacity cost increases.

The models presented in this article build on the well-known *General Optimal Market Area Model* (GOMA) proposed by Erlenkotter.[1] The following section of this article is, therefore, devoted to a brief exposition of the GOMA model along with its roots and extensions. We then extend the model and analysis to the outsourcing and lost sales cases in the next two sections. Concluding remarks and future research directions are provided in the last section.

The General Optimal Market Area Model

Market area models are deeply rooted in the economics literature, starting with central place theories postulated by Lösch[18] and Christaller.[19] A number of market area models have been developed to address important issues such as spatial pricing policies, agglomeration, and land use and these are incorporated in general equilibrium models (Bos[20] and Krugman[21]). Our purpose in using these models is fundamentally different from those of the economists since we are interested in the market areas of facilities that belong to the same firm whereas economists are mostly concerned with the optimal market areas of competing firms with single facilities. Market area models have also been used in operational research, marketing, and regional science. Erlenkotter[1] presents a comprehensive review of this literature.

In operational research, Erlenkotter's GOMA model is arguably the best known optimal market area model. The purpose of GOMA is to determine the size and the market area of one facility in order to minimize the total cost per unit area. Then, the optimal number of facilities can be found by dividing the total area by the optimal market area. GOMA is based on the following assumptions:

(a) demand is uniformly distributed over an infinite plain, with density $D > 0$ per unit area;
(b) the cost of a facility of size w is $Fw^{\alpha} + bw$, where $0 \leq \alpha < 1$, $F > 0$, and $b \geq 0$;
(c) unit transportation costs are related to travel distance r as cr^{β}, where $c > 0$ and $\beta > 0$;
(d) the market area can be circle (C), a hexagon (H), a square (S), or a diamond (D);
(e) distances are measured according to the Euclidean (E) or rectilinear (R) metric.

The market area is given as $A = w/D$. Here and in subsequent analyses, all costs are commensurate, i.e., they apply to the planning horizon. To specify an instance of the model, Erlenkotter uses the notation GOMA(α, β, s, d) in which α and β are the values of exponents given in assumptions (b) and

(c), s denotes one of the market areas shapes in (d), and d denotes one of the distance metrics in (e). If a symbol " * " appears in the position of s or d, the property or the formula is valid for all cases of s or d. Erlenkotter provides a solution for each combination of the above instances. The solution basically depends on what he calls a configuration factor $K(\beta, s, d)$ that reflects the transportation cost exponent, the market area shape, and the distance metric.

We now focus on GOMA(0, 1, *, *) since it will be the basis for the remainder of this article. Note that, $\alpha = 0$ means that the facility cost is a linear function of size and $\beta = 1$ indicates that the transportation cost is linear in the distance traveled. The problem of minimizing the total cost per area is formulated as:

$$\min_{A \geq 0} \mathrm{TC}(A) = \frac{F}{A} + cDK\sqrt{A} \tag{1}$$

where $K\sqrt{A}$ denotes the average distance between a facility and a customer. For the sake of clarity, we suppress the dependence of K on β, s, and d. Note also that Equation (3.3) in Erlenkotter[1] is the total cost per unit demand. Multiplying this equation by the demand density will also yield (1). The variable facility cost b has no impact on TC and is therefore ignored in this section. It is straightforward to prove that the objective function is quasi-convex and if the configuration factor can be obtained in closed form, so can the optimal solution. Hence, the first order condition gives the optimal market area as

$$A^* = \left(\frac{2F}{cDK}\right)^{2/3},$$

and the optimal total cost per unit area as

$$\mathrm{TC}(A^*) = \frac{3}{\sqrt[3]{4}} F^{1/3} (cDK)^{2/3}.$$

It can easily be shown that the optimal cost is quite insensitive to small changes in problem parameters and market area. The values of configuration factor K for the specified market area shapes and distance metrics are given in Table 1. Note that the configuration factor itself is insensitive to different market area shapes. Therefore, as far as the optimal cost is concerned, choosing suboptimal or somewhat irregular market area shape does not result in much inaccuracy. The optimal market area A^*, and thus the number of facilities could, however, be relatively more sensitive to shape.

Earlier optimal market area models, of which GOMA is essentially an extension, are given by Newell[2]

	Distance Metric	
Market Area Shape	Euclidean	Manhattan
Circular	0.376	0.479
Hexagon	0.377	0.480
Square	0.383	0.500
Diamond	0.383	0.471

Table 1: Sample values of $K(1, s, d)$ under different shapes and distance metrics

and Geoffrion.[3] With respect to previous models, GOMA can handle both economies of scale in facility costs and economies of distance in transport costs, as well as different market area shapes and distance metrics. Recently, GOMA has been extended to uneven demand density and cost parameters (Dasci and Verter[22]), to multi-echelon distribution systems (Campbell,[23] Webster and Robinson[24] and Erlebacher and Meller[5]) and refined by incorporating logistics system design components (Erlebacher and Meller[5] and Rutten, van Laarhoven, and Vos[25]).

One study related to ours, by Webster and Gupta,[26] incorporates uncertain and non-stationary temporal demand in GOMA to minimize the expected discounted cost. The authors show that the problem is equivalent to a deterministic case in which demand is equal to the weighted average of expected demand during the planning horizon. However, they do not consider capacity issues. Instead, they implicitly assume that capacity is unlimited or can be expanded at no cost anytime during the planning horizon. Therefore, our setting is fundamentally different from theirs. We consider the case in which the firm has to settle on a fixed capacity before the demand is realized, considering the tradeoffs between excess capacity and potential shortages.

Facility Location and Capacity Acquisition with Outsourcing

We first analyze the outsourcing case. The problem is to determine the optimal market area and the capacity of a facility to minimize the expected total cost, equal to the sum of fixed facility opening costs, variable capacity acquisition and operating costs, transportation costs, and shortage costs. In this case, the shortage cost reflects the price paid to the outside supplier, as well as the additional transportation and handling costs. In what follows, we extend GOMA(0, 1, *, *) to formulate this

problem.

Let $D \geq 0$ denote the random demand per unit area and X denote the capacity acquired at a facility. We assume that demand density is a continuous random variable with a twice-differentiable cumulative distribution function $H(.)$ and a probability density function $h(.)$. The total capacity acquisition and operating cost at a facility is assumed linear, i.e., it is equal to bX, where b denotes the unit variable capacity cost. Any fixed cost related to capacity acquisition is subsumed by the facility fixed cost F. We assume that the total shortage cost is linear at the rate of a. Considering that the event $\{D \leq X/A\}$ corresponds to that of zero shortage, the total shortage cost of a facility is 0 if $D \leq X/A$, and $a(AD-X)$ otherwise. Hence, given A and X, the expected total cost per unit area is a modification of (1) and can be expressed as

$$\begin{aligned} \mathrm{TC}(A,X) &= \frac{F}{A} + \frac{bX}{A} + \int_0^\infty cK\sqrt{A}DdH(D) + \int_{\frac{X}{A}}^\infty \frac{a(AD-X)}{A}dH(D) \\ &= \frac{F}{A} + \frac{bX}{A} + cK\sqrt{A}E[D] + aE[D - X/A]^+ \end{aligned}$$

where $[D - X/A]^+ = \max\{0, D - X/A\}$. Defining the capacity per unit area $y = X/A$, leads to the following reformulation:

$$\mathrm{TC}(A,y) = \frac{F}{A} + by + cK\sqrt{A}E[D] + aE[D-y]^+. \tag{2}$$

Note that the two-dimensional problem $\min_{A,y} \mathrm{TC}(A,y)$ is separable in A and y. The part associated with A is nothing but the GOMA model and hence, it has a unique solution equal to

$$A^* = \left(\frac{2F}{cKE[D]}\right)^{2/3}.$$

Observe that the optimal market area is independent of demand distribution and depends only on the expected demand. It is also independent of capacity acquisition and shortage costs. The other part resembles the well-known newsboy problem and it is straightforward to show that

$$\bar{H}(y^*) = \frac{b}{a},$$

where $\bar{H}(y) = 1 - H(y)$. Hence, the optimal capacity is $X^* = A^*y^*$. No capacity is acquired if outsourcing is cheaper than producing, i.e., $b \geq a$. In this case, the firm satisfies all of the demand from outside.

It is easy to see that the optimal capacity increases as unit shortage cost increases or as unit variable capacity cost decreases. The rate of the change, however, depends on the demand distribution as well as on the size of the optimal market area A^*. The optimal capacity also increases when the spread of the demand distribution increase, while the mean stays constant. Expected demand has mixed effects. A shift in the mean demand increases the optimal capacity per unit area, and hence the total capacity at all facilities, but at the same time it increases the number of facilities. Therefore, the optimal capacity *per facility* may increase or decrease. As a result, the firm may increase capacity either by opening more facilities or by acquiring more capacity per facility or both.

A closed form solution is not available for all distributions. As an illustration, however, let us assume that demand follows an exponential distribution with parameter λ. In this case $E[D] = 1/\lambda$ and the optimal capacity per unit area is

$$y^* = \frac{1}{\lambda} \ln \frac{a}{b}.$$

The optimal market area is

$$A^* = \left(\frac{2\lambda F}{cK}\right)^{2/3},$$

and optimal capacity per facility is then

$$X^* = \left(\frac{2F}{cK\sqrt{\lambda}}\right)^{2/3} \left(\ln \frac{a}{b}\right).$$

Hence, if demand is exponential, the firm opts for more facilities with larger capacities as the expected demand increases, i.e., as λ decreases. In this case, it is also possible to compute the optimal total cost as

$$\mathrm{TC}(A^*, X^*) = \frac{3}{\sqrt[3]{4}} F^{1/3} \left(\frac{cK}{\lambda}\right)^{2/3} + \frac{b}{\lambda}\left(1 + \ln \frac{a}{b}\right).$$

As a second illustration, assume that demand follows a Weibull distribution with scale parameter $1/\lambda$ and shape parameter α. In this case $E[D] = (1/\lambda)\Gamma(1/\alpha + 1)$, where $\Gamma(.)$ is the Gamma function, and the optimal capacity per unit area is

$$y^* = \frac{1}{\lambda}\left(\ln \frac{a}{b}\right)^{\alpha},$$

the optimal market area is

$$A^* = \left(\frac{2\lambda F}{cK\Gamma(1/\alpha + 1)}\right)^{2/3},$$

and the optimal capacity per facility is

$$X^* = \left(\frac{2F}{cK\sqrt{\lambda}\Gamma(1/\alpha+1)}\right)^{2/3}\left(\ln\frac{a}{b}\right)^{\alpha}.$$

Note that expected demand increases as α or λ decreases. Keeping λ constant and assuming that $\ln a/b = 1$, we observe that, as the expected demand increases, i.e., as α decreases, the firm opts for more facilities but with smaller capacities. As these two examples show, capacity per facility does not necessarily increase with mean demand. Sometimes, it can be less costly to set up smaller but more facilities to accommodate increasing demand.

In the remainder of this section we investigate the sensitivity of the optimal cost. It is well known that in GOMA model the optimal cost is insensitive to small changes in the optimal market area and in the problem parameters. In our model, it is even more robust to suboptimal market area choices. Therefore, here we study the impact of suboptimal *capacity* choices on the optimal cost. More precisely, we investigate how the ratio $\mathrm{TC}(A^*, \beta X^*)/\mathrm{TC}(A^*, X^*)$ changes as β changes. However, these results are not parameter-free and we must therefore resort to a numerical study. We have chosen a set of problem parameters such that both parts (i.e., the parts that depend on A and y) of the objective function (2) are close to each other at the optimum. To this end, we set $F = 10,000$, $c = 1$, $K = 0.38$, and $E[D] = 1$. For the pair (a, b), we have chosen five values as (50, 10), (50, 20), (100,10), (100,20), and (100,50). Finally, for the demand distribution we have selected two normal distributions each with mean one and standard deviation of 0.1 and 0.3, and an exponential distribution with parameter one. These three distributions were chosen to control for various degrees of uncertainty. The respective coefficients of variations are 0.1, $0.\bar{3}$, and 1.0, i.e., from a low degree to a higher degree of uncertainty.

The results are summarized in Table 2. The range of β suggests that a suboptimal choice of capacity is set between half and twice the optimal capacity. Each entry in the table is the ratio of the suboptimal cost to the optimal cost. The main observation is that the optimal cost is relatively robust to small capacity changes. However, robustness lessens as unit shortage cost increases or as variable capacity cost decreases, *ceteris paribus*. Similarly, as the coefficient of variation decreases the optimal cost becomes more sensitive. It is therefore particularly important to seek an optimal solution when demand can be estimated with a higher degree of certainty. Also note that it is preferable to invest

	Normal (1, 0.01)					Normal (1, 0.09)					Exponential (1)				
			a/b					a/b					a/b		
	100	50	50	100	100	100	50	50	100	100	100	50	50	100	100
β	50	20	10	20	10	50	20	10	20	10	50	20	10	20	10
0.5	1.28	1.29	1.52	1.77	2.13	1.16	1.18	1.32	1.46	1.68	1.03	1.04	1.09	1.12	1.19
0.6	1.21	1.22	1.38	1.57	1.83	1.11	1.12	1.21	1.30	1.42	1.02	1.03	1.05	1.07	1.11
0.7	1.15	1.15	1.25	1.37	1.52	1.07	1.07	1.12	1.17	1.22	1.01	1.01	1.03	1.04	1.06
0.8	1.08	1.08	1.13	1.19	1.24	1.03	1.03	1.05	1.07	1.09	1.00	1.01	1.01	1.02	1.02
0.9	1.02	1.02	1.03	1.05	1.05	1.01	1.01	1.01	1.02	1.02	1.00	1.00	1.00	1.00	1.01
1.0	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
1.1	1.02	1.02	1.02	1.03	1.02	1.01	1.01	1.01	1.01	1.01	1.00	1.00	1.00	1.00	1.00
1.2	1.08	1.06	1.05	1.07	1.05	1.03	1.03	1.03	1.05	1.04	1.00	1.01	1.01	1.01	1.02
1.3	1.15	1.11	1.08	1.12	1.09	1.07	1.06	1.06	1.09	1.08	1.01	1.01	1.02	1.03	1.04
1.4	1.21	1.16	1.12	1.17	1.12	1.11	1.10	1.09	1.14	1.11	1.02	1.02	1.04	1.05	1.06
1.5	1.28	1.20	1.15	1.22	1.16	1.16	1.14	1.13	1.18	1.15	1.03	1.03	1.05	1.07	1.09
1.6	1.35	1.25	1.18	1.27	1.19	1.22	1.18	1.16	1.23	1.19	1.04	1.04	1.07	1.09	1.12
1.7	1.41	1.30	1.21	1.32	1.22	1.28	1.23	1.20	1.29	1.23	1.05	1.06	1.09	1.12	1.15
1.8	1.48	1.35	1.25	1.37	1.26	1.34	1.27	1.23	1.34	1.26	1.06	1.07	1.12	1.15	1.18
1.9	1.54	1.39	1.28	1.42	1.29	1.40	1.32	1.27	1.39	1.30	1.08	1.09	1.14	1.19	1.22
2.0	1.61	1.44	1.31	1.47	1.33	1.46	1.37	1.31	1.44	1.34	1.09	1.11	1.17	1.22	1.26

Table 2: Sensitivity of the optimal cost with outsourcing.

in excess of the optimal capacity, rather than short of it.

Facility Location and Capacity Acquisition with Lost Sales

We now analyze the case in which shortages become lost sales. The problem is of the same type, as in the outsourcing case, but the transportation cost is somewhat different since if a shortage occurs then there is less to transport. Therefore, even though shortage cost increases, transportation cost decreases with the shortage. This case better reflects reality when outsourcing is impossible in case of a shortage, for example, due to a future industry supply shortage. In this case, the shortage cost reflects the forgone profit and the loss of goodwill. Similar to the earlier model, total shortage cost is assumed to be linear with total shortage. Therefore, the shortage cost has the same expression as in the outsourcing model. The total transportation cost of a facility, however, changes to $cK\sqrt{A}(AD)$ if $D \leq X/A$, and $cK\sqrt{A}(X)$ otherwise. Hence, given A and X, the expected total cost per unit area is:

$$\begin{aligned} \mathrm{TC}(A,X) &= \frac{F}{A} + \frac{bX}{A} + \int_0^{\frac{X}{A}} cK\sqrt{A}D dH(D) + \int_{\frac{X}{A}}^{\infty} (\frac{cK\sqrt{A}X}{A} + \frac{a(AD-X)}{A}) dH(D) \\ &= \frac{F}{A} + \frac{bX}{A} + cK\sqrt{A}E[D] + (a - cK\sqrt{A})E[D - X/A]^+. \end{aligned}$$

Setting again $y = X/A$, the total cost can be expressed as

$$\mathrm{TC}(A,y) = \frac{F}{A} + by + cK\sqrt{A}E[D] + (a - cK\sqrt{A})E[D-y]^+.$$

The two dimensional problem $\min_{A,y} \mathrm{TC}(A,y)$ can be solved sequentially: $\min_y \mathrm{TC}(y)$, where $\mathrm{TC}(y) = \min_A \mathrm{TC}(A,y)$. Let $A^*(y)$ denote the optimal market area for a fixed y, then the problem can be re-stated as

$$\min_{y \geq 0} \mathrm{TC}(y) = \left\{ by + aE[D-y]^+ + \min_{A(y) \geq 0} \left\{ \frac{F}{A(y)} + cK\sqrt{A(y)}(E[D] - E[D-y]^+) \right\} \right\}. \qquad (3)$$

The minimization in the brackets is nothing but the GOMA model and hence, it has a unique solution given as

$$A^*(y) = \left(\frac{2F}{cK(E[D] - E[D-y]^+)} \right)^{2/3}.$$

Substituting back into (3), the problem becomes

$$\min_y \mathrm{TC}(y) = \left\{ by + aE[D-y]^+ + c'(E[D] - E[D-y]^+)^{2/3} \right\},$$

where $c' = 3/\sqrt[3]{4}F^{1/3}(cK)^{2/3}$. Note that $\mathrm{TC}(y)$ is continuous, $\mathrm{TC}(0) = aE[D]$, and $\mathrm{TC}(y) \to \infty$ as $y \to \infty$. Therefore, there exists a finite optimal solution. Unfortunately, $\mathrm{TC}(y)$ does not lead to a closed form solution and worse, it is neither convex nor quasi-convex. Nonetheless, we prove that $\mathrm{TC}(y)$ has either two or no local extrema for distributions that have increasing failure rates (i.e., $h(D)/\bar{H}(D)$ is increasing in D). This is not unduly restrictive since important and practical classes of distributions, including exponential, normal, and Weibull with shape parameter greater than one, have increasing failure rates. To this end, let $\mathrm{TC}'(y)$ and $\mathrm{TC}''(y)$ denote the first and second derivatives of $\mathrm{TC}(y)$ with respect to y. $\mathrm{TC}'(y)$ is derived as:

$$\mathrm{TC}'(y) = b - a\bar{H}(y) + \frac{2c'\bar{H}(y)}{(E[D] - E[D-y]^+)^{1/3}}.$$

Lemma 1 *$TC'(y) \to \infty$ as $y \to 0$ and $TC'(y) \to b$ as $y \to \infty$.*

Proof: Straightforward.

Lemma 2 *If D has increasing failure rate then $TC''(y) = 0$ has at most one solution.*

Proof: Compute

$$\frac{d^2\mathrm{TC}(y)}{dy^2} = h(y)\left[a - \frac{2c'}{3(E[D] - E[D-y]^+)^{1/3}}\right] - \frac{2c'\bar{H}(y)^2}{9(E[D] - E[D-y]^+)^{4/3}}.$$

Rearranging the terms for $\mathrm{TC}''(y) = 0$ yields

$$\frac{9a}{2c'} = \frac{1}{(E[D] - E[D-y]^+)^{1/3}}\left[3 + \frac{\bar{H}(y)}{h(y)}\frac{\bar{H}(y)}{(E[D] - E[D-y]^+)}\right].$$

The left-hand side is constant. We show that the right-hand side is monotonically decreasing in y, so that there can be at most one solution. Note that $d(E[D] - E[D-y]^+)/dy = \bar{H}(y) \geq 0$ and $d\bar{H}(y)/dy = -h(y) \leq 0$ for all $y \geq 0$. Therefore, the first term on the right-hand side and the last term in brackets are decreasing. Since $\bar{H}(y)/h(y)$ is nothing but the reciprocal of the failure rate, it is also decreasing with y. Hence, the right-hand side is decreasing in y. □

Combining the lemmas, it is clear that $\mathrm{TC}'(y)$ is quasi-convex and there can be at most two values where it vanishes. We have identified three possible cases in which above conditions are satisfied. Figure 1 depicts these cases. In case (a), the total cost has a local maximum and a local minimum and therefore, the optimal solution is either at the local minimum or at $y = 0$. In the other two cases, (b) and (c), the optimal solution is at $y = 0$, which means that the firm should not acquire any capacity. In these cases, A^* approaches infinity, which translates into no facility opening. There is also a fourth case in which $\mathrm{TC}'(y) = 0$ has only one solution. In this case, $\mathrm{TC}(y)$ has a saddle point but it is similar to case (b).

It is relatively easy to characterize these cases and find the optimal solution numerically. First, one should check whether $\mathrm{TC}'(y)$ approaches b from above or from below. If it does from above, then we have case (c). If it does from below, then the local minimum of $\mathrm{TC}'(y)$, which then exists, should be found. If $\mathrm{TC}'(y)$ is positive at the minimum, then we have case (b). If it is negative, then we

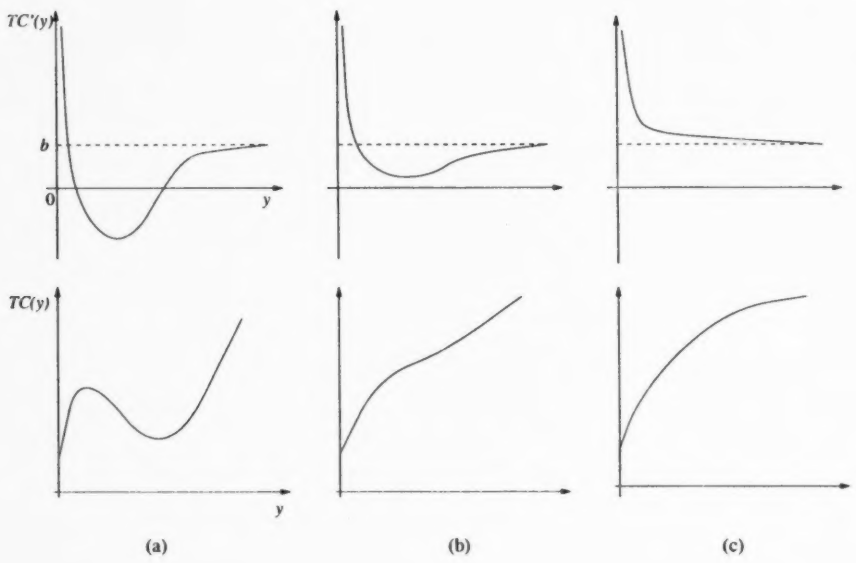

Figure 1: Three possible cases for $TC'(y)$ and $TC(y)$.

have case (a) and one should search for a zero of $TC'(y)$, which is greater than its minimum point. This point will be the local minimum of $TC(y)$. As a last step, the total cost at the local minimum of $TC(y)$ should be compared to $TC(0) = aE[D]$. If the demand distribution does not possess the increasing failure rate property a more involved numerical search is required, which is not a major problem considering that this is a single variable equation.

It is easy to see that the optimal market area in the lost sales case is *always* larger than the optimal market area in the case with outsourcing, regardless of demand distribution and other parameters. In other words, more facilities are opened when outsourcing is used. This difference is due to the fact that in the outsourcing case the firm ships all of the demand, whereas in the lost sales case some of the potential demand might not be satisfied and hence not shipped. Therefore, in the outsourcing case a higher expected transportation activity leads to the opening of more facilities which in turn reduces the overall transportation cost. A similar conclusion for the respective capacities is neither forthcoming nor needed because all of the parameters play a complex role. In particular, shortage cost a has different meanings and hence, different values in these cases.

In the remainder of this section we analyze the sensitivity of the optimal cost. We use the same set

	Normal (1, 0.01)					Normal (1, 0.09)					Exponential (1)				
			a/b					a/b					a/b		
	100	50	50	100	100	100	50	50	100	100	100	50	50	100	100
β	50	20	10	20	10	50	20	10	20	10	50	20	10	20	10
0.5	1.19	1.14	1.31	1.62	1.94	1.11	1.07	1.20	1.38	1.57	1.01	*	1.05	1.09	1.16
0.6	1.15	1.11	1.24	1.46	1.69	1.07	1.05	1.13	1.25	1.36	1.01	*	1.03	1.06	1.09
0.7	1.10	1.08	1.16	1.31	1.44	1.04	1.03	1.08	1.14	1.19	1.00	*	1.02	1.03	1.05
0.8	1.06	1.04	1.09	1.16	1.21	1.02	1.02	1.03	1.06	1.08	1.00	*	1.01	1.01	1.02
0.9	1.02	1.01	1.02	1.04	1.05	1.01	1.00	1.01	1.01	1.02	1.00	*	1.00	1.00	1.00
1.0	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	*	1.00	1.00	1.00
1.1	1.02	1.02	1.02	1.03	1.02	1.01	1.00	1.01	1.01	1.01	1.00	*	1.00	1.00	1.00
1.2	1.07	1.05	1.05	1.07	1.05	1.02	1.02	1.03	1.04	1.04	1.00	*	1.01	1.01	1.01
1.3	1.14	1.10	1.08	1.12	1.09	1.05	1.04	1.05	1.08	1.07	1.00	*	1.01	1.02	1.03
1.4	1.20	1.14	1.11	1.17	1.12	1.09	1.07	1.08	1.13	1.11	1.01	*	1.02	1.04	1.05
1.5	1.27	1.19	1.14	1.22	1.15	1.14	1.10	1.11	1.18	1.15	1.01	*	1.03	1.06	1.08
1.6	1.33	1.24	1.18	1.27	1.19	1.19	1.14	1.15	1.23	1.18	1.02	*	1.05	1.08	1.11
1.7	1.40	1.28	1.21	1.32	1.22	1.24	1.18	1.18	1.28	1.22	1.02	*	1.06	1.11	1.14
1.8	1.47	1.33	1.24	1.36	1.26	1.30	1.22	1.21	1.32	1.26	1.03	*	1.08	1.13	1.17
1.9	1.53	1.37	1.27	1.41	1.29	1.35	1.26	1.25	1.37	1.30	1.04	*	1.09	1.16	1.20
2.0	1.60	1.42	1.31	1.46	1.32	1.41	1.30	1.28	1.42	1.33	1.04	*	1.11	1.19	1.24

Table 3: Sensitivity of the optimal cost with lost sales.

of parameters as in the previous section. The results are summarized in Table 3. In one of the cases, denoted with a column of asterisks, the optimal solution corresponds to zero capacity. We therefore exclude it from further consideration. The general observation is similar to the outsourcing case; the optimal cost is relatively robust to small changes in capacity and the robustness diminishes as shortage cost increases or as capacity cost decreases. Similarly, as the coefficient of variation decreases the optimal cost becomes more sensitive. However, the optimal cost in this model is more robust than the previous model for all sets of parameters.

Concluding Remarks

We have studied the facility location and capacity acquisition problem under demand uncertainty. Our analytical approach has provided simple models for a potentially difficult problem. Shifting the focus from a precise representation of facility location to the market areas of facilities has led to simple algebraic expressions that are either solvable in closed form or amenable to efficient solution procedures. We have derived a number of interesting insights into the system dynamics, which would

take laborious numerical studies with a conventional mathematical programming approach. Clearly, these two approaches do not compete but rather complement each other. An analytical approach would provide managers with valuable insights about the structural decisions and possible alternatives, whereas a more detailed discrete model, based on those insights, could prescribe optimal configurations.

We have observed that the issue of how the firm deals with shortages has important implications in a spatial environment. We have considered two possibilities: In the first one, shortages are satisfied by purchasing from outside and in the second one, they become lost sales. In a non-spatial environment, this distinction is not important. However, in our setting they have led to two different models, in which location and capacity decisions may differ substantially. We have shown that fewer facilities are opened in the lost sales case than the outsourcing case, regardless of the underlying parameters. These two variants have also led to substantially different formulations. While the outsourcing case is solvable in closed form or in implicit form, the lost sales case has proved more challenging in that it can only be tackled numerically. However, the numerical requirement is nothing more involved than a search for the minimum of a single-variable continuous function.

We believe that the models introduced here open new avenues for further research in facility location and capacity planning. Some of the possible extensions can easily be accommodated by our models. For example, more general forms of capacity cost, including any type of piecewise linear function, can be analyzed in a similar manner. Solving the problem for each technology can also incorporate technology selection decisions. Discrete demand distribution and discrete capacity options may even lead to simpler models. Throughout the text, however, we have implicitly assumed that the uncertainty unfolds in the same manner everywhere in the spatial market. Even though this is a plausible assumption for a simplified analysis, modeling local variations in demand realization should lead to an important extension. Finally, a market area model that considers dynamic capacity planning or multi-echelon production distribution system should constitute an important contribution in this field.

Acknowledgements: This work was partially supported by the Canada Research Chair in Distribution Management and by the Canadian Natural Sciences and Engineering Research Council under grant number OGP0039682. This support is gratefully acknowledged.

References

1. D. Erlenkotter, "The general optimal market area model," *Annals of Operations Research*, vol. 18, pp. 45–70, 1989.

2. G.F. Newell, "Scheduling, location, transportation, and continuum mechanics: Some simple approximations to optimization problems," *SIAM Journal on Applied Mathematics*, vol. 25, pp. 346–360, 1973.

3. A.M. Geoffrion, "The purpose of mathematical programming is insight, not numbers," *Interfaces*, vol. 7, no. 1, pp. 81–92, 1976.

4. A.M. Geoffrion, "Making better use of optimization capability in distribution system planning," *AIIE Transactions*, vol. 11, pp. 96–108, 1979.

5. S.J. Erlebacher and R.D. Meller, "The interaction of location and inventory in designing distribution systems," *IIE Transactions*, vol. 32, pp. 155–166, 2000.

6. A.S. Manne, *Investment for Capacity Expansion: Size, Location, and Time-Phasing*, MIT press, Cambridge, MA, 1967.

7. D. Erlenkotter, "Capacity planning for large multilocation systems: Approximate and incomplete dynamic programming," *Management Science*, vol. 22, pp. 274–285, 1975.

8. R.C. Rao and D.P. Rutenberg, "Multilocation plant sizing and timing," *Management Science*, vol. 23, pp. 1187–1198, 1977.

9. C.O. Fong and V. Srinivasan, "The multiregion dynamic capacity expansion problem, Part I," *Operations Research*, vol. 29, pp. 787–799, 1981a.

10. C.O. Fong and V. Srinivasan, "The multiregion dynamic capacity expansion problem, Part II," *Operations Research*, vol. 29, pp. 800–816, 1981b.

11. C.O. Fong and V. Srinivasan, "The multiregion dynamic capacity expansion problem: An improved heuristic," *Management Science*, vol. 32, pp. 1140–1152, 1986.

12. J.G. Klincewicz, H. Luss, and C.-S. Yu, "A large-scale multilocation capacity planning model," *European Journal of Operational Research*, vol. 34, pp. 178–190, 1988.

13. S.-K. Lim and Y.-D. Kim, "An integrated approach to dynamic plant location and capacity planning," *Journal of the Operational Research Society*, vol. 50, pp. 1205–1216, 1999.

14. S. Ahmed and R. Garcia, "Dynamic capacity acquisition and assignment under uncertainty," Tech. Rep., The Logistics Institute, Georgia Institute of Technology, Atlanta, GA, 2001.

15. V. Balachandran and S. Jain, "Optimal facility location under random demand with general cost structure," *Naval Research Logistics Quarterly*, vol. 23, pp. 421–436, 1976.

16. L.J. LeBlanc, "A heuristic approach for large scale discrete stochastic transportation-location problem," *Computers and Mathematics with Applications*, vol. 3, pp. 87–94, 1977.

17. G. Laporte, F.V. Louveaux, and L. Van hamme, "Exact solution to a location problem with stochastic demands," *Transportation Science*, vol. 28, pp. 95–103, 1994.

18. A. Lösch, *The Economics of Location*, Yale University Press, New Haven, CT, 1954.

19. W. Christaller, *Central Places in Southern Germany*, Prentice-Hall, Englewood Cliffs, NJ, 1966.

20. H.C. Bos, *Spatial Dispersion of Economic Activity*, Rotterdam University Press, 1965.

21. P. Krugman, "Space: The final frontier," *Journal of Economic Perspectives*, vol. 12, pp. 161–174, 1998.

22. A. Dasci and V. Verter, "A continuous model for production-distribution system design problem," *European Journal of Operational Research*, vol. 129, pp. 287–298, 2001.

23. J.F. Campbell, "One-to-many distribution with transshipments: An analytic model," *Transportation Science*, vol. 27, pp. 330–340, 1993.

24. S. Webster and Jr. E.P. Robinson, "Analytical results for two-echelon distribution network design," *Location Science*, vol. 2, pp. 223–239, 1994.

25. W.G.M.M. Rutten, P.J.M. van Laarhoven, and B. Vos, "An extension of the GOMA model for determining the optimal number of depots," *IIE Transactions*, vol. 33, pp. 1031–1036, 2001.

26. S. Webster and A. Gupta, "The general optimal market area model with uncertain and non-stationary demand," *Location Science*, vol. 3, pp. 25–38, 1995.